

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,300

Open access books available

130,000

International authors and editors

155M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



A Model of the Continual Adaptive Online Knowledge Assessment System

Miran Zlatović, Igor Balaban and Željko Hutinski

Abstract

This chapter presents a model of a novel adaptive online knowledge assessment system and tests the efficiency of its implementation. System enables continual and cumulative knowledge assessment, comprised of sequence of at least two interconnected assessments, carried-out throughout a reasonably long period of time. Important characteristics of the system are: (a) introduction of new course topics in every subsequent assessment, (b) re-assessment of earlier course topics in every subsequent assessment iteration, (c) in an adaptive manner, based on student's achievements during previous assessments. Personalized post-assessment feedback guides each student in preparations for upcoming assessments. The efficiency has been tested on a sample of 78 students. Results indicate that the proposed adaptive system is efficient on an individual learning goal level.

Keywords: online knowledge assessment, adaptive knowledge assessment, improving classroom teaching, post-secondary education, learning strategies, learning goals

1. Introduction

The courses taught by the authors of this chapter (ICT-oriented, undergraduate university level courses) use a type of accumulative model of tracking students' activities, where multiple traditional written mid-term assessments grant most of the points required to pass the course. A more specific feature of this tracking model is that the units of learning contents are assessed multiple times. In other words, every subsequent mid-term assessment includes the re-assessment of previous content too, but with diminishing contributions – for example, 2nd mid-term assessment might include 40% of the content from the 1st mid-term assessment and 60% of new content, 3rd mid-term might include 10% of the oldest content (1st mid-term), 30% of the older content (2nd mid-term) and 60% of brand new content, etc.

Although we were generally satisfied (in terms of overall course grades) with the results of our traditional non-adaptive pen-and-paper assessment approach, we wanted to explore the possibilities of including Information and Communications Technology (ICT) support and adaptive assessment into the accumulative tracking model, to achieve following improvements:

1. To adapt the re-assessment portion of the mid-term to each individual student, based on the results he/she obtained for that content during previous mid-term:
 - Students that have shown higher levels of mastery of particular content during previous mid-term need not be re-assessed about that content in detail – i.e. they may receive less questions or less complex questions (to demonstrate that they have not forgotten what they had known before).
 - Students that have shown lower levels of mastery of old content should be re-assessed about that content more thoroughly (to demonstrate that currently they have more knowledge than before).
2. To include ICT support into adaptive knowledge assessment process, because manual adaptation of subsequent pen-and-paper mid-terms for each individual student, as suggested above, would be too complex to manage.

Such assessment model should be **continual** (span across multiple connected assessments throughout semester) and **adaptive** (re-assessment part of every subsequent assessment would be adapted to each participant, based on his/her previous results). It would also need to fit our current teaching delivery model, i.e. traditional classroom courses supported by blended e-learning and activity tracking.

With that respect, our primary goal was to explore the possibilities of improving the in-house knowledge assessment process by making it adaptive, but without introducing the complexity of complete adaptive learning systems (which will be mentioned briefly in the opening paragraphs of Chapter 3).

2. Research methodology

To achieve the desired goal (as mentioned in Introduction - inclusion of ICT support and adaptive assessment into our existing accumulative tracking model), we used Design and Development Research (DDR) Method that allows researchers to establish new procedures, techniques and tools based on specific needs analysis [1] and that consists of seven iterative phases [2]. Within the first, “Focus” phase we bounded the scope of the project to ensure that the project pursues an important goal that can be achieved with current resources, which is presented in the introduction section. Within the “Understand” phase we analyzed research literature to investigate the problem (section “The Context of the Study”). Research objectives and hypothesis were then identified within the “Define” phase. The initial solution was designed under the “Conceive” phase (section “Rationale Behind the Proposed Model”). The “Build” phase aimed at developing the model and building a test platform (section “Development of the Model”). We evaluated the efficacy and behavior of the solution in a real context within the “Test” phase (section “Testing the Model”). This chapter, in overall, is part of the last phase in the DDR methodology (“Present”), where we elaborate how the developed solution contributed to solving the problem.

3. The context of the study

The adaptive online education is highly represented in current scientific and professional research, especially the studies focused on adaptive learning and adaptive learning systems (ALS). Here we refer to adaptive learning as a process which creates unique learning experiences for every learner by taking into consideration many learner’s

traits, such as his/her interests, performance, personality, etc. [3]. Most research efforts in the ALS field are focused on full e-learning systems, which are driven by two main principles: (a) selection and delivery of the appropriate learning contents to each participant, so that (b) each participant can improve the effects of his/her education [4, 5].

Although this chapter follows a similar principle, it does not focus on adaptive education in its broader and general sense. Instead, it puts an emphasis on the process of adaptive online knowledge assessment [6–8], i.e. on the process of selection and application of different types of questions within written online knowledge assessment, in order to improve each student’s achievement levels of learning goals. In the context of this chapter, like the approach taken by the Stanford University, we consider learning goals as the statements of “... what we want our students to be able to demonstrate at the end of our class.” [9]. Examples of such learning goals can be found in **Table 1**. Achievement of such learning goals can be measured by standard knowledge assessment grading techniques.

3.1 Adaptation and learning strategies vs. learning styles

To be able to consider users’ individual differences, ALSs rely on user models [10] that keep track of many elements, including learning styles, learners’ personal preferences, prior knowledge, skills and competences. Many studies stress the importance of learning styles during adaptation process. As shown by Soflano, Connoly and Hainey [11], the adaptation based on learning styles in games-based learning (GBL) environment allowed learners to complete the tasks faster, compared to both non-adaptive GBL and to classic textbook learning. Tseng, Chu,

Online tests	Learning goal codes	Learning goal descriptions
First test (t1)	LG1	Define decision support systems and expert systems
	LG2	Describe the elements, components, objectives and functions of IS
	LG3	Describe the structure of decision support systems and expert systems
	LG4	Describe the decision-making process and the role of DSS and ES in the decision-making process
	LG5	Describe the types of IS
	LG6	Distinguish the life and development cycle of IS
	LG7	Define types of content search on the Internet
	LG8	Define common Internet services
	LG9	Describe the elements of a computer network
	LG10	Describe the elements, functions and structure of the Internet, intranets and extranets
	LG11	Describe and compare the types of content search on the Internet
	LG12	Describe the ISO/OSI model and TCP/IP model
	LG13	Describe common Internet services
Second test (t2)	LG14	Define concepts of multimedia and virtual reality
	LG15	Describe multimedia systems and virtual reality
Third test (t3)	LG16	Define concepts of the safety and security of IS
	LG17	Describe and explain the safety and security of IS

Table 1.
Learning goals used in Adaptivity application to test the model.

Hwang and Tsai [12] report that the approach based upon multiple sources of personalization (learning behavior and personal learning style) is helpful in improving both the learning achievements and learning efficiency of individual students.

Although adaptations based on learning styles might have a role in improving learning achievements when applied on an entire ALS level, it should be noted that learning styles would not be that useful if they were used as a foundation for adaptations within narrower field of knowledge assessment only. Hartley [13] claims that individual learning styles are mostly static in time and not easily changed, unlike learning strategies which are primarily dynamic, conditioned by current tasks and can be manipulated with during shorter periods. Hartley defines learning strategies as "... the different combinations of activities (i.e. 'strategies') students use while learning.". Similarly, Mayer [14] defines them as "... behaviors of a learner that are intended to influence how the learner processes information". For this chapter, we consider these two basic learning strategies - deep and surface learning, which can be briefly described as follows [15, 16]:

- Surface learning – any combination of the activities used by the students while learning, that lead to the learning aimed at mere reproduction of the contents. Understanding of learning contents is very low or non-existent.
- Deep learning – any combination of the activities used by the students while learning, that lead to the learning aimed at understanding of the contents, i.e. questioning of alternatives, raising additional questions, etc.

We aim to use the feedback part of proposed adaptive system to steer the students towards behaviors and activities which would preferably lead to deep learning while preparing for the re-assessment of earlier learning contents.

Learning strategies, as described above, can be measured by various instruments, such as Study Process Questionnaire [17], although their direct measurement is not in the scope of this chapter. Here we refer to additional study which has shown that learning strategies can be facilitated (stimulated) and have important influence on achievement levels of learning goals [18] – it has been shown that an announcement of any type (form) of online knowledge assessment is not suitable for the facilitation of the more desirable deep learning and that all learning strategies facilitated by such announcements do not equally contribute to the achievement levels of the required learning goals. Zlatović, Balaban and Kermek [18] have demonstrated that a deep learning strategy has a positive effect on results in both essay and multiple-choice types of online assessment, while surface learning strategy has a negative impact on results in online essay, and no impact on results in online multiple-choice question assessment. When it comes to the levels of knowledge, the study has demonstrated that achievements of lower levels of knowledge (rote memorizing, reproduction, understanding) have been primarily stipulated by surface learning strategies which were facilitated by using online assessments containing multiple-choice questions. Achievements of higher levels of knowledge (analysis, synthesis and evaluation) were better when essay-based online assessments were used to facilitate deep learning strategies. Due to all these findings, we decided to incorporate the effects of learning strategies facilitation in proposed model, as an important supportive element in the adaptation of the re-assessment of the old learning contents.

3.2 Feedback

Another major aspect of our model involves feedback which is a major element of quality in teaching and assessment [19, 20]. Students also appreciate the value of

feedback and are aware of its importance in achieving learning goals [21]. Maier, Wolf and Randler [22] have examined feedback effects with computer-assisted multiple-tier tests and it was revealed that feedback is more effective when it is designed as elaborated (specific) feedback and that the elaborated feedback is effective when it is perceived as helpful.

By using feedback based on the results of individual's current assessment, the adaptive system we propose will announce to each individual student the following instructions related to the re-assessment part of the next assessment:

1. What type of questions is predominantly going to be used in the following iteration (e.g. essay-oriented questions, matching the terms, fill-in the blanks, multiple choice) and
2. What is the expected difficulty of those questions (i.e. easy, medium or hard)?

Such announcements are supposed to facilitate the appropriate learning strategies (preferably deep learning) during preparations for re-assessment.

3.3 Adaptation throughout a series of assessments

The central aspects of the model we are proposing are the **continuity** of the assessment and the **adaptation** between the series of the connected assessments (i.e. the adaptive re-assessment part of each subsequent assessment).

Review in the field of the adaptive online knowledge assessment reveals that historically most efforts are focused on studying various aspects of adaptability within a single knowledge assessment, usually within a self-assessment and/or formative assessment [23, 24].

However, to continuously monitor students' progress, a continuous knowledge assessment was proposed. McAlpine [25] defines it as "... the more modern form of modular assessment, where judgments are made at the end of each field of study". Continuous knowledge assessment belongs to the group of formative assessment techniques, since it provides plethora of individuals' learning progress indicators while students are still committed to the learning process. Therefore, such indicators can be used to carry-on corrective actions while the teaching process is still ongoing – e.g. to adapt teaching process to the specific needs of participants.

Continuous formative evaluation using ALS system Amrita Learning [26] uses multiple assessments in adaptive manner, but each assessment covers different learning contents and old contents are never re-assessed. Therefore, such adaptation process does not consider the results of earlier assessment(s).

Grundspenkis [8] and Grundspenkis and Anohina [27] have described an adaptive learning and assessment system where concept maps are used as a more machine-friendly replacement for essays. Course contents are introduced gradually in time, through multiple stages. Every subsequent stage can only upgrade existing content from previous stage with new concepts. Adaptive knowledge assessments take place between stages, but although these assessments encompass contents from all available stages (similarity with our approach), the adaptivity is still limited to a single assessment. Adaptivity is reflected via two properties: (i) student can request a task with reduced difficulty, if initial version is too difficult and (ii) system can automatically increase the difficulty of the following task if the student has achieved required score without any reductions. Still, there is no evidence that e.g. an assessment that takes place between stages 2 and 3 takes into consideration the results from the assessment conducted between stages 1 and 2. There are also examples of adaptive and continuous assessment within commercial e-learning

platforms – e.g. Khan Academy, whose approaches towards assessing students' mastery of a particular topic is described in [28]. Historically, the Khan Academy used the streak concept, where student had to solve correctly at least 10 problems in a row. Then the system assumes that required proficiency level has been achieved and student can progress further to new topics. More advanced proficiency model replaced the streak approach – next task was selected using logistic regression techniques, considering both previously solved tasks and current proficiency level of a student. While the element of adaptivity over the series of assessments is present, it still lacks the systematic inclusion of older content into upcoming assessments.

Within the area of ALSs we often encounter distinctions mentioning micro- and macro-adaptation. It is suggested by Van Lehn [29] that primary focus of macro-adaptation is application of adaptivity on a global task selection process within entire ITS, while primary focus of micro-adaptations are lower-level in-task interactions. Knowledge assessment is usually considered to belong to the micro-level of an ITS. Results of assessments are then used to update learner models, which are then used in subsequent macro-adaptation activities [24, 30]. Since we propose the adaptive model of continual assessment that is designed primarily to be used standalone, without being part of a larger ALS or ITS, macro-level of adaptation will be represented by adapting the re-assessment part of the next assessment. Results of micro-activities (individual assessments) would update simplified user model (user's achievement levels per topic/learning goal), which is later used to perform macro-adaptation between two assessments.

Review of the available research suggests that sufficient investigation effort has not yet been put into assessment systems which implement adaptivity within series of interconnected assessments, specifically into systems using adaptivity to re-assess previous learning contents. Additional insights about such systems is one of the scientific contributions of this study.

4. Research objectives and hypothesis

In respect to the issues noted from the research literature, the objectives of this study are as follows:

To develop and test a model of the adaptive online knowledge assessment system that facilitates those learning strategies that lead towards better achievement of the required learning goals.

To provide feedback to students based on individual results of their online assessments, containing suggestions about the assessment types that are going to be used for the re-assessment of particular learning goals in the next iteration of an adaptive assessment.

In line with the research objectives, the following hypothesis is formulated:

The model is proposed, of the continual adaptive online knowledge assessment system, which leads to better achievements of the required levels of learning goals, by utilizing a personalized feedback to announce what questions types will be selected in the following assessment iteration and by utilizing learning strategies facilitated by such personalized feedback.

5. Rationale behind the proposed model

Based on the findings and the experience from previous research regarding the learning strategies, as well as the other relevant work indicated in previous section, we propose the model of an adaptive online knowledge assessment system, which

supports series of assessments connected in a linear way, in a chain-like structure. It is designed to guide the individual towards continuous improvements in achievement levels of required learning goals within traditional higher education class-based courses by focusing on several key aspects:

Assessment process is carried out continually during longer period (e.g. one semester), throughout a series of assessments following the principles generally common for ALS described in previous section. Necessity for having longer period is also supported by findings from Dembo and Praks-Seli [31], stating that changes in students' learning strategies cannot appear instantaneously, due to them (strategies) being either part of individual's automated behavior patterns or being carried over from other courses.

Personalized feedback per assessed learning goal (for example, see Maier, Wolf and Randler [22]) will be presented at the end of each assessment, based on individuals' achievement levels per topic/learning goal, suggesting what type of questions will be used next time, to re-assess those learning goals.

Given the application of above-mentioned feedback and having enough time between two assessments, individuals have enough time to adjust their learning strategies [18, 31] – preferably towards deep learning, so that they are more likely to improve their achievement levels in re-assessed topics.

Inclusion of the following aspects into the proposed assessment model is part of the original contribution of this chapter:

- Every subsequent assessment includes re-assessment of the topics from previous assessments (continual assessment of topics, to stimulate improvements of learning goals' achievement levels – here we build upon findings from the field of cognitive psychology, where it was shown "... that repeated testing of information produces superior retention relative to repeated study, especially when testing is spaced out over time." [32]).
- Adaptive re-assessment of old topics based on individuals' previous achievement levels per old topic/learning goal.

The knowledge assessment model proposed in this chapter represents a type of **continual** (carried-out through multiple iterations during longer period of time, i.e. one semester) and **cumulative** (iterations cannot be considered as mutually independent, because subsequent iterations include **earlier content** alongside **newly introduced content**) knowledge assessment. The first iteration (first assessment) is always non-adaptive. Adaptive assessment phase starts with the second iteration of e-assessment by analyzing individual assessment results from the first iteration, which opens-up a possibility to personalize each students' questions structure just for the re-assessment part of the old topics. In those phases system automatically selects the questions (their number, type and difficulty), based on the built-in adaptivity rules which consider student's previous level of learning goals achievements for a particular learning object (topic). At the end of each assessment, system presents the student with the feedback containing information about the level of achievement per learning goal and the types of questions that will be preferred in upcoming assessment to re-assess earlier learning content (especially for those units of content whose learning goals were not met in a satisfactory manner in current assessment). This information should incite students to change learning strategies they intend to use for the re-assessment of earlier learning content.

Inclusion of adaptivity elements within the above-described type of assessment, as well as modeling and development of a system which selects the types of questions to facilitate learning strategies, which in turn lead to a better achievement of the required learning goals, is an important contribution of this chapter.

6. Development of the model

6.1 Basic structure

Following general practices from the field of adaptive knowledge assessment are integrated within the proposed model (references to the numberings 1 to 3 will be used later in the text as “general practice 1”, “general practice 2” and “general practice 3”):

1. Quantitative expression of individuals’ success in achieving particular learning goal [34], e.g., using percentage scale that mimics grading system.
2. Multi-level qualitative marking of questions difficulty, in context of the assessment of associated learning goal, e.g. easy/medium/difficult [33, 34].
3. Rule-based approach towards adaptation process [33, 35].

Besides those elements, continual and cumulation properties are paired with adaptivity features are also built into the model. Cumulation property enables the inclusion of desired elements of adaptivity in the assessment system, in a sense that re-assessment of the earlier learning content may become individualized and in accordance with the achievements examinees have demonstrated during previous iterations:

- Individual goal achievements from the previous iterations can be used to formulate the announcement of the type and the difficulty of the assessment that will be used to re-assess these goals in a new iteration.
- The system informs each examinee what type of the assessment will be used in re-assessment of various portions of earlier learning content, so that (i)
 - Such announcements provide **individual** facilitation of learning strategies, and
 - Effects of the facilitated learning strategies lead to the improvements of students’ performance.

The basic structure of the proposed assessment model is shown in **Figure 1**. The **cognitive level** is a label assigned to a learning goal, according to Bloom’s Taxonomy [36]: 1 – Knowledge ... 6 – Evaluation. It is used to classify learning goals regarding their cognitive levels.

The **learning objects** represent broader units of learning content, to which one or more learning goals are connected.

A **learning goal** is always connected to a particular learning object and a particular cognitive level is assigned to it. Goals also have defined percentage-based thresholds for achievement levels. If the achievement level is below the lowest level, it means that the related learning goal is not achieved; gradual increase in the thresholds reached represents the achievement on a gradually higher level.

The **questions** element represents the assessment questions database and “general practice 2” was followed here. Each question is assigned to one or more **learning goals**. Model supports various types of questions: (i) multiple-choice questions (both single- and multiple correct answers), (ii) matching questions, (iii) fill-in the blanks and (iv) essay questions. Difficulty of a question within the context of particular learning goal [33] is defined by attaching mandatory qualitative label to each question – three levels of difficulty are supported: easy questions (DL1, “difficulty level 1”), medium-difficulty questions (DL2) and difficult questions (DL3).

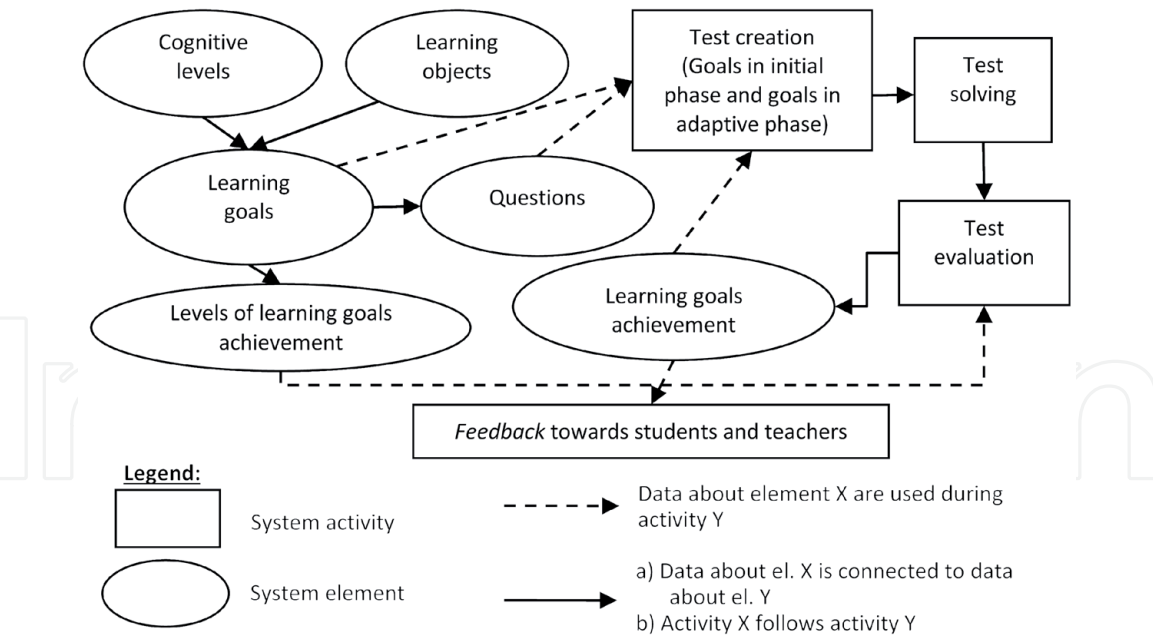


Figure 1.
Basic elements of the proposed model of the continual adaptive online knowledge assessment system.

All the above-mentioned elements (cognitive levels, learning objects, learning goals, question difficulty levels) are defined manually by the teacher within the proposed system – it is solely their responsibility to set-up the database of interrelated learning goals, objects and questions.

The **assessment creation** activity is a central element of the system and takes into consideration all the other main elements of the system, except for feedback, and also leans on general practice (general practice 3). Learning goals that are being assessed for the first time during an assessment cycle are in the initial phase, which means that adaptivity rules do not apply yet. The goals that are re-assessed in the following iterations are in the adaptive phase and the process of questions selection is fully governed by the adaptive rules and results achieved for that goal in previous iteration.

The **learning goals achievement** element is calculated during the assessment evaluation activity, in-line with the “general practice 1”. It is a quantitative indicator of student’s level of achievement of a learning goal, expressed as a percentage scale. Although arbitrary number of thresholds can be used to express various achievement levels, proposed model is set to mimic the traditional grading scale:

- Fail (F or 1): 0–49,99%
- Sufficient (D or 2): 50–62,49%
- Good (C or 3): 62,5–74,99%
- Very good (B or 4): 75–87,49%
- Excellent (A or 5): 87,5–100%

The **feedback towards the students** (see Table 2 for an example) visualizes the individual achievement levels related to the particular learning goals included in assessment and provides personalized suggestions describing what type and difficulty of questions will be used predominantly in following adaptive iteration, during repeated assessment of old learning content.

(1) Results per Learning Goal (LGs)	(2) Announcement of question types and difficulties to be used for a learning goal re-assessment:
LG: Describing the decision process and the role of DSS and ES within it. No. of questions: 3 Max. points: 3 Points achieved: 3 (100%) Learning goal achievement level: Excellent	If the current achievement level of a learning goal is “Fail”: - > it will be re-tested using difficult questions (predominantly using more demanding essay-type questions) If the current achievement level of a learning goal is „Sufficient “or „Good“: - > it will be re-tested using medium and difficult questions (predominantly using essays and matching terms/statements questions; less likely by fill-in-the-blanks and multiple-choice questions) If the current achievement level of a learning goal is “Very good” or “Excellent”: - > it will be re-tested using easy and medium difficulty questions (predominantly using multiple choice, fill-in-the-blanks and matching questions; less likely by short and less demanding essay questions)
LG: Describing the network topologies and elements used to build computer networks No. of questions: 4 Max. points: 5 Points achieved: 3.33 (66.6%) Learning goal achievement level: Good	
LG: Describing the ISO/OSI model and TCP/IP model No. of questions: 3 Max. points: 5 Points achieved: 0 (0%) Learning goal achievement level: Fail	
etc. results for other learning goals	

Table 2.
Excerpt from an automated feedback presented to student at the end of each assessment.

6.2 Flow of the assessment

The first assessment iteration in the assessment cycle is always non-adaptive, as illustrated in **Figure 2**. In this iteration, since it is the first time that all topics are being assessed, all students will have identical structure of the test. Only teacher (without intervention of the built-in adaptivity mechanics) decides (a) which learning objects and goals to include, (b) what difficulty levels of the questions will be required to assess particular learning goal and (c) how many questions (of required difficulty and type) will be included in the test. Besides already mentioned criteria (objects/goals, difficulty and number of questions), teacher can also define that in the initial phase of the assessment all student will be given either: (i) fully identical set of questions, or (ii) randomly selected questions, or (iii) a mixture of fixed and randomly selected questions.

Based on individual results from the first iteration, it is possible to adaptively automate and personalize each student’s questions structure for the re-assessment

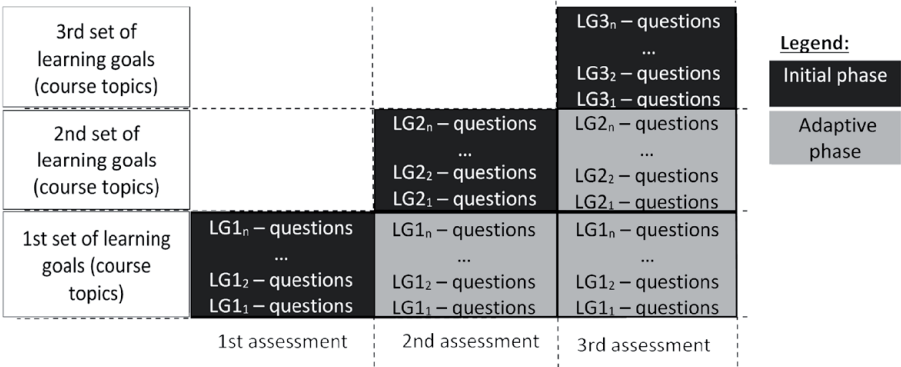


Figure 2.
Flow of the continual and cumulative adaptive knowledge assessment.

of old learning goals in the following iteration. Therefore, the second (and each subsequent) iteration of the assessment implements the cumulation property and it is comprised of the:

- First assessment of **new learning objects** – since these objects enter the assessment for the first time (dark gray rectangles in **Figure 2**), teacher is again responsible for defining all the parameters (as described above), and
- Repeated assessment of **learning objects from the previous iteration** (Property of cumulation) – since these objects need to be assessed repeatedly, they enter the adaptive assessment phase (light gray squares in **Figure 2**) and only the system automatically selects the questions (their number, type and difficulty), based on built-in adaptivity rules (general practice 3) which consider student's previous level of learning goals achievements for that object. Teacher does not have any influence on the question selection process for learning goals that are being assessed repeatedly.

Likewise, the N-th iteration is also cumulative in nature – it includes the first assessment of new learning objects (initial phase with identical assessment structure for all students, teacher defines all parameters for question selection) and the repeated assessment of learning objects which were included in all the previous iterations (without teacher's influence, governed only by built-in adaptivity rules, General practice 3).

Automated process of selecting the questions for learning goals that have entered the adaptive phase relies on five adaptive rules, which will be briefly summarized in following section, for the completeness and clarity of the chapter. More elaborate descriptions and case studies of those rules can be found in [37].

6.3 Adaptive rules

There are three categories of adaptive rules used to select questions for learning goals which have reached the adaptive phase of the assessment. Rules are built around general practices (general practices 1 and 2) and the properties of continuity and cumulation:

1. **Three rules (R1 to R3) to decide the questions difficulty** – the difficulty selection is based on the individual student's achievement level for a learning goal (i.e. score for a group of questions pertaining to that learning goal), in a way that if the achievement level in the previous iteration was:
 - “**Fail**”: select only high-difficulty questions (i.e. highest difficulty available) for that learning goal. This is **rule R1** with the rationale: “improve the non-satisfactory achievement level”.
 - “**Sufficient**” or “**Good**”: select medium- and high-difficulty questions available for that learning goal. This is **rule R2** with the rationale: “maintain decent achievement level, with incentive for improvement”.
 - “**Very good**” or “**Excellent**”: select easy and medium-difficulty questions available for that learning goal: This is **rule R3** with the rationale: “don't forget about this portion of learning content”.
2. **Rule R4 which decreases the number of questions used in the adaptive phase** – both repeated assessment of learning goals from previous iterations

and inclusion of the first-time assessed new learning goals lead to inevitable question inflation (i.e. ever increasing number of questions) and consequently to assessment duration issues (i.e. ever longer duration of the test, to compensate for the ever increasing number of questions). If **N** question were used in **1st iteration**, then at most **N/2** questions will be used in **2nd iteration** for that learning goal, at most **N/3** in **3rd iteration**, etc.

3. Rule R5 which increases the number of questions only for the individuals with low achievement – this rule is complementary to the rule R4. If some student has achieved the lowest (i.e. “Fail”) level for some learning goal during previous iteration, then due to this rule system will individually increase (only for such student) the total number of questions used to re-assess only that failing learning goal. Rule R5 uses the amount obtained from rule R4 as baseline and adds to it. Nevertheless, it also ensures that the total amount of questions for the re-assessment of failed goal does not exceed the number **N** (no. of questions used for that goal in the first iteration). The rationale behind this rule is the following: because of the student’s previous poor achievement for a learning goal, its re-assessment in current adaptive phase should be more thorough for such student.

Regarding rule R1, at first it may seem pedagogically wrong to use only the difficult questions during the re-assessment of failed learning goals. It may very well be perceived as a punishment, but only if those difficult questions **actually were more difficult** than all the questions used in the previous iteration for that learning goal. The responsibility to avoid such unwanted situation lays on the teacher – he/she must include an appropriate mixture of easy, medium and difficult questions for the initial stage of each learning goal. In such circumstances, rule R1 **cannot select even more difficult questions** for failed goals during the adaptive phases – it will merely focus on the pool of questions marked as “difficult” (from the same pool which has already been used in the first iteration), while disregarding less difficult questions. And according to the rules R4 and R5, re-assessment of failed goal also includes less questions, albeit all of them being marked as “difficult”.

7. Testing the model

Adaptivity, the web application for continual adaptive online knowledge assessment, was developed based on the proposed model and built upon Microsoft ASP.NET platform (MS Windows Server, MS SQL Server and ASP.NET) in order to test the model. However, detailed description of the web application is not in the scope of this chapter. More elaborate description of Adaptivity’s architecture can be found in Zlatović and Balaban [38].

The procedure of testing the effectiveness of the model involved approximately half of the students who regularly attended classes at the “Informatics 2” (convenience sample, $N = 78$), which is held at authors’ university as a part of the undergraduate curriculum for the bachelor’s degree in the field of information systems and technology. All students enrolled in “Informatics 2” were divided into two groups (alphabetically, by Faculty administration). We selected randomly one of those groups to participate in experiment. The course is elective and is being taught at the bachelor university level, with first-year students being enrolled predominantly (more than 90% of the population). It is also available for students who attend 2nd and 3rd year of the bachelor program.

Formal curriculum of the course prescribed four written assessments (hereinafter tests) during the semester. The first test was used to verify the functionality of the proposed system in a real environment and under the workload generated by the actual number of users. Therefore, the three remaining tests were included in the research. The type of assessment was cumulative, meaning that each subsequent test included new learning materials along with the old one (as illustrated in **Figure 2**). With respect to the terminology used in previous Section, the individual test in the experimental group matches one iteration within the proposed model of the assessment. All the tests were conducted in strictly controlled environment (in Faculty's computer labs, under teachers' supervision).

7.1 Changes in learning goals achieved

In this section we analyze the results achieved by using Adaptivity to explore whether its usage increased levels of achievement of learning goals that had not been considered satisfactory in previous iterations. **Table 1** shows all the learning goals (LGs) that were examined during the three tests cycle (tests t1, t2 and t3).

Upon completion of all three tests, average achievement scores per learning goals were compared. Prior to any comparisons, all individual achievement scores were converted from absolute points into relative percentages. Absolute points would not make sense here, because each student's assessment in adaptive phase will have different amount of questions used (due to built-in adaptive rules R4 and R5 in particular) and consequentially, absolute points maximum would differ from student to student. Although the distribution of the achievement scores of learning goals in all three iterations did not follow normal distribution (both Kolmogorov–Smirnov and Shapiro–Wilk tests were used), the size of the experimental group ($N = 78$) is large enough to warrant the usage of parametric t-tests [39]. Specifically, two-tailed paired samples t-tests were conducted, because pre- and post-test scores produced by the same students were compared.

Table 3 shows the results of the comparisons made at the end of each iteration and **Table 4** the results of the comparisons made between the first and the final test. Only the learning goals which elicited significant increase or decrease in the average achievements score were kept in those tables. In the first cycle, learning goals LG14 and LG15 were not calculated, because in the second test (t2) those goals were assessed for the first time, so there were no results for them from the previous iteration. Likewise, when displaying the results of the second cycle, LG16 and LG17 are not shown either. Item pairs in tables are encoded using simple LGx_ty scheme, where LGx stands for Learning Goal X ($1 \leq x \leq 17$) and ty stands for particular test iteration y ($1 \leq y \leq 3$) – e.g. LG6_t2 represents the score of Learning Goal 6 in test iteration 2.

Paired-samples t-test statistics from **Table 3** show that at the end of the 1st cycle of assessment, only 4 learning goals displayed significant changes in average achievement scores – for three of them (LG8, LG9 and LG13) there is significant increase of the average scores (ranging from 6.51% to 12.18% higher score on the average), while one learning goal (LG10) displayed significant decrease of the average score (17.37% lower score on the average). After the 2nd cycle, statistically significant increases of the average scores were noted for 6 learning goals in total (LG1, LG6, LG7, LG9, LG10 and LG13, ranging from 6.98% to 12.95% higher score on the average) and one learning goal (LG15) has shown statistically significant decrease of the average score (9.14% lower score on the average).

After the 2nd cycle, LGs from 1 to 13 have been adaptively re-tested for the second time, while LGs 14 and 15 have been adaptively re-tested for the first time. Lack of the statistically significant difference in score for LG8 after the 2nd cycle

Differences after the 1st test cycle (test t2 vs. test t1)	Paired Diff. Mean	t	Sig. (2-tailed)
LG1_t2 - LG1_t1	-1.28205	-0.217	0.829
LG6_t2 - LG6_t1	6.19692	1.462	0.148
LG7_t2 - LG7_t1	1.06410	0.183	0.855
LG8_t2 - LG8_t1	6.51282	2.501	0.014
LG9_t2 - LG9_t1	9.38013	3.267	0.002
LG10_t2 - LG10_t1	-17.37179	-3.479	0.001
LG11_t2 - LG11_t1	3.57859	1.035	0.304
LG13_t2 - LG13_t1	12.17949	2.838	0.006
Differences after the 2nd test cycle (test t3 vs. test t2)	Paired Diff. Mean	t	Sig. (2-tailed)
LG1_t3 - LG1_t2	10.25641	2.432	0.017
LG6_t3 - LG6_t2	9.88244	2.232	0.029
LG7_t3 - LG7_t2	10.89744	2.194	0.031
LG8_t3 - LG8_t2	1.91026	0.716	0.476
LG9_t3 - LG9_t2	6.97538	2.074	0.041
LG10_t3 - LG10_t2	12.94872	2.406	0.019
LG11_t3 - LG11_t2	6.10051	1.750	0.084
LG13_t3 - LG13_t2	7.13141	2.236	0.028
LG14_t3 - LG14_t2	6.72962	1.808	0.075
LG15_t3 - LG15_t2	-9.13500	-2.462	0.016

Table 3.
Comparison of the achievements of learning goals between consecutive tests - comparisons after the 1st cycle (test t2 vs. test t1) and the 2nd cycle (test t3 vs. test t2) of the continual assessment, paired samples t-test (N = 78, df = 77, p < 0.05).

Differences between the first and the final test (test t3 vs test t1)	Paired Diff. Mean	t	Sig. (2-tailed)
LG1_t3 - LG1_t1	8.97436	1.867	0.066
LG6_t3 - LG6_t1	16.07936	4.090	0.000
LG7_t3 - LG7_t1	11.96154	2.256	0.027
LG8_t3 - LG8_t1	8.42308	2.443	0.017
LG9_t3 - LG9_t1	16.35551	5.341	0.000
LG10_t3 - LG10_t1	-4.42308	-0.991	0.325
LG11_t3 - LG11_t1	9.67910	2.468	0.016
LG13_t3 - LG13_t1	19.31090	4.437	0.000

Table 4.
Comparison of the achievements of learning goals between the final and the first test - comparisons between the final test (t3) and the first test (t1) of the continual assessment, paired samples t-test (N = 78, df = 77, p < 0.05).

can be interpreted as the stagnation (compared to the significant increase LG8 has had after the 1st cycle) – slight average increase of 1.91% cannot be taken as statistically significant at $p < 0.05$. Differences in achievement levels for learning goal LG11 show stagnation after both 1st and 2nd cycle of the assessment. Interestingly,

Table 4 suggests that LG11 has significantly higher average score when entire chain of the assessments is taken into consideration.

Results shown in **Table 4** (final test t3 vs. first test t1) include only those learning goals that have been used throughout entire chain of assessments, i.e. only LGs from 1 to 13 (LGs 14 and 15 were introduced in test t2 for the first time, while LGs 16 and 17 were introduced in test t3 for the first time). At the end of the series of assessments, 6 learning goals in total (LG6, LG7, LG8, LG9, LG11 and LG13) have shown statistically significant increase of the average achievement score (ranging from 8.42% to 19.31% on the average).

The results of one learning goal (LG10) have effectively canceled themselves out during the repeated assessments – data from **Table 3** shows that LG10 recorded significant decrease of the score after the 1st cycle and significant increase of the score after 2nd cycle – the results for LG10 after test t3 have become similar to the initial results after test t1. This is shown as statistically insignificant decrease of 4.42% on the average in **Table 4**. Although final results for LG10 indicate stagnation, initial significant decrease of students' score after LG10's first re-assessment has been compensated by significant increase after the second re-assessment of LG10. Similar reasoning can be applied to LG1 too – the decrease of the score after the 1st cycle was not large enough to be considered significant and the increase of the score after the 2nd cycle was significant (**Table 3**). But the final results for LG1 (in **Table 4**) suggest that observed increase for LG1 between the last (3rd test) and the first assessment (1st test) is borderline insignificant at $p < 0.05$, because students had achieved slightly lower score at LG1 during 3rd assessment than during 2nd.

In addition to the already discussed LG10 and LG1, for 5 more learning goals in total (LG2, LG3, LG4, LG5 and LG12) repeated assessment did not cause statistically significant changes in average scores and those LG's were omitted from **Tables 3** and **4**. These results can also be interpreted as the stagnation in the achievement levels.

Based on those indicators, it is shown that the use of the proposed model encourages improvements in the level of achievement for almost 50% of the evaluated learning goals (6 out of 13 goals which have been included in the assessment from the beginning), or at least it enables the retention of the existing levels of the achievement (7 out of 13 goals which have been included in the assessment from the beginning). Constant decrease of the achievement levels has not been noticed at any of the learning goals which have been re-assessed at least twice.

8. Discussion

It has been demonstrated that the application of the Model has positive influence on improving achieved levels of knowledge per individual learning goals being assessed. During the three-test assessment cycle, it was shown that for 6 learning goals there was a global tendency of improving the achievement (i.e constantly increased achievement levels during re-assessments of those learning goals) - predominantly for the more complex goals, which required the ability to describe and understand concepts, not just to recall the facts. For 5 learning goals, there was a global tendency to maintain previous level of achievement. Only one learning goal showed negative initial result, although, as already described, after 2nd iteration that learning goal recorded significant improvement in scores, but not adequate to globally overcome the low score after the 1st iteration. And the improvements for one more learning goal were borderline insignificant.

It has been mentioned in Section 7 that only half of the student population enrolled in course "Informatics 2" were used to test the model (i.e. "experimental group"). One could ask why the results obtained during model testing have not been

compared with the results of the other half of the class (i.e. “control group”). Main reason is that there have been too many differences in the overall knowledge assessment process between two groups, for the comparisons to be valid and meaningful. While the “experimental” half of the class used online Adaptivity system, which had provided mixture of various types of questions (multi-choice, fill-in, match, essay), between-assessment adaptation and individualized post-assessment feedback per learning goal, students in so-called “control” half of the class were given only pen-and-paper tests using essay-type questions exclusively, without detailed feedback and without any form of adaptation (i.e. the traditional way of administering the summative assessments within the course).

It must be mentioned that number of re-assessments per LO and LG used in this research (one initial assessment and at most two adaptive re-assessments) may not be enough in terms of proper continual knowledge assessment. Since the assessment results of the experimental group had to be used as a formally valid substitute for the final summative results of the “Informatics 2”, the assessment process design for the experimental group could not have diverged too far from the assessment process used for the rest of the class. E.g. fixed and relatively small number of assessments per semester was one of the constraints that had to be adhered to. It would be highly recommended to use more frequent (re)assessments in future research. Nevertheless, despite relatively low number of re-assessments, the proposed model did yield at least the retention of the previously reached levels of achievements (for 7 of 13 LGs), if not slight improvements in levels of achievements during re-assessments (for 6 of 13 LGs).

Another valid question is what type of knowledge has been taught and the type of teaching used. Content of the “Informatics 2” course is related to purely theoretical knowledge, within the area of expertise in ICT belonging to both social and technical sciences. Teaching process had consisted of purely ex-cathedra lectures with supplementary slides and lectures available within learning management system (LMS). Because of the assessed knowledge nature, success percentages in the Adaptivity have been set to mimic traditional grading system, requiring at least 50% success for a positive grade. If necessary, grading scales in Adaptivity can be re-adjusted to fit other areas of expertise, where higher cut-off points may be required for positive grades.

Most of the LGs (see **Table 1**) used in this study are focused on lower levels of knowledge. While not ideal, it is consistent with findings in [8] that even the most sophisticated automated assessment systems do not allow for testing of knowledge which is higher than level 3 or 4 in Bloom’s taxonomy. Adaptivity as a system does support usage of essay-type of questions, which must be graded manually by teachers. Therefore, higher levels of knowledge could also be re-assessed in the continual adaptive manner, at the expense of re-introducing increased teachers’ workload.

Overall, those findings are in-line with traditional features of continuous assessment, i.e. the ability to apply corrective actions while the education is still ongoing [25, 40] and the superior retention of information due to repeated testing spaced-out over time [32]. These are also in-line with several observations given in [41]: (i) assessment should not encourage surface learning and (ii) adaptive assessment provides benefits to both summative and formative assessment.

Application of the proposed system also helps alleviate one of the biggest practical disadvantages of manual continuous assessment reported in literature – vastly increased teachers’ workload, due having to spend more time to prepare and carry out frequent activities to track their learners [30, 42]. Proposed system is fully automating the adaptive portion of the continual re-assessment of old topics, leaving the teacher with task to manually create only the content related to the new topics, which are being assessed for the first-time.

Thus, it is shown that the Model, which employs continual and cumulative approach towards knowledge assessment and which: (a) individually adjusts amount, difficulty and type of questions per learning goal, based on previously demonstrated levels of achievement of learning goals, and (b) announces what types of the assessment will be used to test particular learning goals in the upcoming iteration, has predominantly positive effects on individual's success at the level of particular learning goals, therefore supporting research objectives and hypothesis.

9. Research limitations and future research suggestions

This research was conducted among ICT-oriented higher education students, which have already been using online education before. Therefore the sample used may not represent well the population from other fields of higher education (natural, technical, biomedical, humanistic, etc.) or outside of the higher education (e.g. secondary education, workplace education and/or life-long learning, etc.). Inclusion of respondents from other areas would ensure more varied population of respondents. Also, research was conducted within a course that uses blended education model (mixture of traditional class-based education and elements of online education), therefore it is advised to exercise caution when trying to generalize the results of this study to institutions and environments that practice either self-paced education, full online education, or traditional class-based education. The specifics of the assessment process itself represent another limitation – the assessment was adjusted to fit the continuous monitoring of students' activities in the context of high education that adheres to Bologna Process.

The course was taught by the authors themselves and the authors have also designed the assessments, so a methodological bias needs to be considered when analyzing the results of this study. Further research should include both courses taught by and assessments designed by other teachers too.

We have also included only learner's cognitive abilities. Affective characteristics of students (e.g., motivation, mastery goal orientation), which can also be important when designing adaptive assessment system, were not included. Further research should include broader student modeling. In line with [40], further research could also expand onto teacher responsiveness, which builds upon continuous results provided by the proposed assessment system.

On a different note, the current implementation of the Model could be a worthy contribution to further development of the Adaptive Learning Management systems that consider various users' individual differences. Integration of the proposed Model in such adaptive environment as a complementary to the adaptive lessons could present a significant step forward in the design and implementation of Adaptive Learning Management systems.

10. Conclusion

This study describes original approach related to the modeling and implementation of the continual adaptive online knowledge assessment within class-based courses, where the adaptive aspects of assessment are used to re-assess old topics and are:

1. Applied within the series (or chains) of assessments, and
2. Based on the results that students have achieved in the previous assessments, rather than being based on the results achieved in the current, isolated assessment.

The Model introduces adaptation throughout a series of assessments in order to continuously monitor students and uses immediate feedback (mostly based on recommendations from Rowe and Wood [21] and Maier, Wolf and Randler [22]) as a major element of quality in teaching and assessment, which is given to students at the end of each assessment to facilitate the appropriate learning strategies.

The empirical study of the Model's efficiency has shown that it is possible to design the system for adaptive online knowledge assessment, which can facilitate desirable learning strategies, which in turn lead to the achievement of required learning goals by announcing and using the appropriate types of questions in assessments.


Since it was shown that continual and cumulative adaptive online assessment is an efficient tool for facilitation of the appropriate learning strategies, the results of this chapter can be useful to the educational institutions when designing and implementing online knowledge assessments within class-based courses. The proposed Model also fits particularly well in continual monitoring and evaluation of students' activities which is in line with Bologna Process, and in the same time relieves teachers from heavier workload.

Author details

Miran Zlatović, Igor Balaban* and Željko Hutinski
University of Zagreb, FOI, Varaždin, Croatia

*Address all correspondence to: igor.balaban@foi.hr

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Richey RC, Klein, JD. Design and development research. New Jersey: Lawrence Erlbaum Associates Inc.; 2007.
- [2] Easterday MW, Rees Lewis DG, Gerber EM. The logic of design research. *Learning: Research and Practice*. 2017;1-30. DOI: 10.1080/23735082.2017.1286367
- [3] Yaghmaie M, Bahreininejad A. A context-aware adaptive learning system using agents. *Expert Systems with Application*. 2011;38(4):3280-3286. DOI: 10.1016/j.eswa.2010.08.113
- [4] Hafidi M, Bensebaa T, Trigano P. Developing adaptive intelligent tutoring system based on item response theory and metrics. *International Journal of Advanced Science and Technology*. 2012;43:1-14
- [5] Ahuja NJ, Sille R. A critical review of development of intelligent tutoring systems: Retrospect, present and prospect. *International Journal of Computer Science Issues*. 2013;10(2):39-48
- [6] Avgerinos E, Karageorgiadis A. On an adaptive formative assessment platform for STEM Education. In: *IEEE Global Engineering Education Conference (EDUCON)*; 2017. IEEE; 2017. p. 1216-1224. DOI: 10.1109/EDUCON.2017.7943003
- [7] Comas-Lopez M, de la Rubia MA, Sacha GM. Adaptive test system for subjects that simultaneously include theoretical content and numerical problem solving. In: *International Symposium on Computers in Education (SIIE)*; 2018; Jerez, Spain. IEEE; 2018. p. 1-5. DOI: 10.1109/SIIE.2018.8586729
- [8] Grundspenkis J. Intelligent Knowledge Assessment Systems: Myth or Reality. In: Lupeikiene A et al., editors. *Frontiers in Artificial Intelligence and Applications*, vol. 315, Databases and Information Systems X. IOS Press; 2019. p. 31-46. DOI: 10.3233/978-1-61499-941-6-31
- [9] Learning Goals. Teaching Commons, Stanford University [Internet]. Available from: <https://teachingcommons.stanford.edu/resources/course-preparation/creating-syllabus/learning-goals> [Accessed 2020-06-25]
- [10] Normadhi NBA, Shuib L, Nasir HNM, Bimba A, Idris N, Balakrishnan V. Identification of personal traits in adaptive learning environment: Systematic literature review. *Computers & Education*. 2019;130:168-190. DOI: 10.1016/j.compedu.2018.11.005
- [11] Soflano M, Connolly TM, Hainey T. An application of adaptive games-based learning based on learning style to teach SQL. *Computers & Education*. 2015;86:192-211. DOI: 10.1016/j.compedu.2015.03.015
- [12] Tseng JCR, Chu H-C, Hwang GJ, Tsai C-C. Development of an adaptive learning system with two sources of personalization information. *Computers & Education*. 2008;51(2):776-786. DOI: 10.1016/j.compedu.2007.08.002
- [13] Hartley J. *Learning and studying: a research perspective*. USA: Routledge; 1998. DOI: 10.4324/9780203132951
- [14] Mayer R. Learning strategies: An overview. In: Weinstein C et al., editors. *Learning and Study Strategies: Issues in Assessment, Instruction, and Evaluation*. New York: Academic Press; 1988. p. 11.22. DOI: 10.1016/B978-0-12-742460-6.50008-6
- [15] Hoeksema LH. *Learning Strategy as a Guide to Career Success in Organizations*. Leiden University, Netherlands: DSWO Press. 1995.

- [16] Marton F, Säljö R. On qualitative differences in learning: I - Outcome and process. *British Journal of Educational Psychology*. 1976;46:4-11. DOI: 10.1111/j.2044-8279.1976.tb02980.x
- [17] Biggs JB, Kember D, Leung DYP. The revised two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*. 2001;71:133-149. DOI: 10.1348/000709901158433
- [18] Zlatović M, Balaban I, Kermek D. Using Online Assessments to Stimulate Learning Strategies and Achievement of Learning Goals. *Computers & Education*. 2015;91:32-45. DOI: 10.1016/j.compedu.2015.09.012
- [19] Peacock S, Murray S, Kelly J, Scott A. The transformative role of ePortfolios: Feedback in healthcare learning. *International Journal of ePortfolio*. 2011;1(1):33-48
- [20] Pacheco-Venegas ND, López G, Andrade-Aréchiga M. Conceptualization, development and implementation of a web-based system for automatic evaluation of mathematical expressions. *Computers & Education*. 2015;88:15-28. DOI: 10.1016/j.compedu.2015.03.021
- [21] Rowe AD, Wood, LN. Student perceptions and preferences for feedback. *Asian Social Science*. 2008;4(3):78-88. DOI: 10.5539/ass.v4n3p78
- [22] Maier U, Wolf N, Randler C. Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*. 2016;95(1):85-98. DOI: 10.1016/j.compedu.2015.12.002
- [23] Conejo R, Guzmán E, Trella M. The SIETTE automatic assessment environment. *International Journal of Artificial Intelligence in Education*. 2016;26(1):270-292. DOI: 10.1007/s40593-015-0078-4
- [24] Chrysafiadi K, Troussas C, Virvou M. A framework for creating automated online adaptive tests using multiple-criteria decision analysis. In: *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*; 2018; Miyazaki. IEEE; 2018. p. 226-231. DOI: 10.1109/SMC.2018.00049
- [25] McAlpine M. *Principles of assessment*. Luton: CAA Centre, University of Luton; 2002.
- [26] Raman R, Nedungadi P. Adaptive learning methodologies to support reforms in continuous formative evaluation. In: *IEEE International Conference on Educational and Information Technology (ICEIT)*; 2010. IEEE; 2010. vol. 2, p. V2-429. DOI: 10.1109/ICEIT.2010.5607608
- [27] Grundspenkis J, Anohina A. Evolution of the concept map based adaptive knowledge assessment system: Implementation and evaluation results. *Scientific Journal of Riga Technical University. Computer Sciences*. 2009;38(38):13-24. DOI: 10.2478/v10143-009-0001-2
- [28] Hu D. How Khan Academy is using Machine Learning to Assess Student Mastery [Internet]. 2011. Available from: <http://david-hu.com/2011/11/02/how-khan-academy-is-using-machine-learning-to-assess-student-mastery.html> [Accessed 2020-06-25]
- [29] VanLehn K. The behavior of tutoring systems. *International journal of artificial intelligence in education*. 2006;16(3):227-265.
- [30] Rus V, Baggett W, Gire E, Franceschetti D, Conley M, Graesser A. Towards Learner Models based on Learning Progressions (LPs) in DeepTutor. In: Sottolare RA et al., editors. *Design Recommendations for*

Intelligent Tutoring Systems, Vol 1. U.S. Army Research Laboratory; 2013. p. 183-192.

[31] Dembo MH, Praks Seli H. Students' Resistance to Change in Learning Strategies Courses. *Journal of Developmental Education*. 2004;27(3): 2-11.

[32] Larsen DP, Butler AC, Roediger III HL. Test-enhanced learning in medical education. *Medical Education*. 2008;42(10):959-966. DOI: 10.1111/j.1365-2923.2008.03124.x

[33] Hatzilygeroudis I, Koutsojannis C, Papachristou N. Adding adaptive assessment capabilities to an e-learning system. In: Mylonas P et al, editors. *SMAP '06, First International Workshop on Semantic Media Adaptation and Personalization*; 2006; Athens. IEEE; 2006. p. 68-73. DOI: 10.1109/SMAP.2006.8

[34] Chatzopoulou DI, Economides AA. Adaptive assessment of student's knowledge in programming courses. *Journal of Computer Assisted Learning*. Wiley. 2010;26(4):258-269. DOI: 10.1111/j.1365-2729.2010.00363.x

[35] Jadhav M, Rizwan S, Nehete A. User profiling based Adaptive Test Generation and Assessment in E-Learning System. In: *IEEE 3rd International Advance Computing Conference (IACC)*; 2013. IEEE; 2013. p. 1425-1430. DOI: 10.1109/IAdCC.2013.6514436

[36] Bloom BS, Engelhart MD, Furst EJ, Hill W, Krathwohl DR. *Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook I: Cognitive Domain*. USA: McKay Press; 1956.

[37] Zlatović M, Balaban I. Personalizing Questions Using Adaptive Online Knowledge Assessment. In: *eLearning 2015 - 6th International Conference on*

e-Learning; 2015; Belgrade. Univerzitet Metropolitan: Belgrade; 2015. p. 185-190.

[38] Zlatović M, Balaban I. Adaptivity: A Continual Adaptive Online Knowledge Assessment System. In: Rocha Á et al, editors. *Trends and Innovations in Information Systems and Technologies; WorldCIST 2020; Advances in Intelligent Systems and Computing*, vol 1161. Cham: Springer. 2020. p. 152-161. DOI: 10.1007/978-3-030-45697-9_15

[39] Lumley T, Diehr P, Emerson S, Chen L. The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual Review of Public Health*. 2002;23:151-169. DOI: 10.1146/annurev.publhealth.23.100901.140546

[40] Preciado-Babb P, Metz M, Sabbaghan S, Davis B. The Role of Continuous Assessment and Effective Teacher Response in Engaging all students. In: Hunter R, editor. *Mathematical Discourse that Breaks Barriers and Creates Space for Marginalized Learners*. Leiden, The Netherlands: Brill|Sense; 2018. p. 101-119. DOI: 10.1163/9789463512541_006

[41] Challis D. Committing to quality learning through adaptive online assessment. *Assessment & Evaluation in Higher Education*. 2005;30(5):519-527. DOI: 10.1080/02602930500187030

[42] Timmerman K, Doom T. Infrastructure for Continuous Assessment of Retained Relevant Knowledge. *ACM Inroads*. 2017;8(2):73-77. DOI: 10.1145/3095781.3017738